Molecular Phylogenetics and Evolution 61 (2011) 149–156

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Reconciling gene trees of eastern North American minnows

Phillip R. Hollingsworth Jr. *, C. Darrin Hulsey

Department of Ecology and Evolutionary Biology, University of Tennessee, 569 Dabney Hall, Knoxville, TN 37996, USA

ARTICLE INFO

Article history:
Received 8 February 2011
Revised 27 May 2011
Accepted 30 May 2011
Available online 6 June 2011

Keywords:
Cyprinidae
Phylogenetic incongruence
Species tree
*BEAST
Benthic/limnetic habitat

ABSTRACT

Most eastern North American cyprinid fishes belong to a clade known as the "open posterior myodome" (OPM) minnows, but phylogenetic relationships within this clade have been difficult to ascertain. Previous attempts to resolve relationships among the generally benthic "chubs" and the more pelagic "shiners", that constitute the majority of OPM minnows, have led to highly discordant phylogenetic hypotheses. To further examine relationships among the OPM minnows, we utilized both a concatenated Bayesian approach and a coalescent-based species tree method to analyze data from six protein coding nuclear loci (Enc1, Ptr, Ryr3, Sh3px3, Tbr1, and Zic1), as well as the mitochondrial locus (Cytb). We focused our analyses on the chub-like genus *Phenacobius*, a group that has drifted topologically between other benthic chubs and the more pelagic shiners, and also included exemplar taxa from 11 other OPM lineages. Individual gene trees were highly discordant regarding relationships within *Phenacobius* and across the OPM clade. The concatenated Bayesian analysis and coalescent-based species tree reconstruction recovered slightly different phylogenetic topologies. Additionally, the posterior support values for clades using the coalescent-based approach were consistently lower than the concatenated analysis. However, *Phenacobius* was resolved as monophyletic and as the sister lineage to *Erimystax* regardless of the combined data approach taken. Furthermore, *Phenacobius* + *Erimystax* was recovered as more closely related to the shiners we examined than to other chubs. Relationships within *Phenacobius* varied depending on the combined phylogenetic method utilized. Our results highlight the importance of multi-locus, coalescent-based approaches for resolving the phylogeny of diverse clades like the eastern North American OPM minnows.

1. Introduction

Fishes in the family Cyprinidae dominate the freshwater habitats of North America (NA) with over 300 species distributed from Canada south to the Neovolcanic Plateau in southern Mexico (Burr and Mayden, 1992). The eastern half of NA contains a particularly high number of cyprinid lineages, and the vast majority comprise a single clade united by the morphological synapomorphy of an open posterior myodome (OPM) (Mayden, 1989; Simons and Mayden, 1997; Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010). It has been hypothesized based on osteological characters that there are two major monophyletic groups within the hyperdiverse OPM clade: (1) the generally benthic "chubs" and (2) the more pelagic "shiners" (Mayden, 1989). However, in subsequent phylogenetic studies, relationships among constituent chub and shiner lineages have varied considerably (Coburn and Cavender, 1992; Simons and Mayden, 1997; Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). In order to clarify the relationships among major lineages of minnows in eastern NA, we analyzed sequence data of six nuclear DNA (nDNA) loci and one mitochondrial DNA (mtDNA) locus using exemplars of recognized genera within the OPM clade. We also included complete taxon sampling of species in the genus *Phenacobius* in order to examine relationships within this genus as well as the relationship of *Phenacobius* to other OPM genera. *Phenacobius* was placed in the chub clade by Mayden (1989) and Coburn and Cavender (1992) based on osteology, but recent molecular phylogenies suggest it may be more closely related to the shiner OPM minnows (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). Therefore, determining the phylogenetic position of this genus of five species poses an interesting problem and focal point for this multi-locus phylogenetic analysis of OPM minnow relationships.

Phylogenetic trees reconstructed from individual loci do not necessarily reflect the species tree of a given clade because discordance among gene trees may be common, especially in clades characterized by large population sizes and short intervals between diversification events (Degnan and Rosenberg, 2009). Such

* Corresponding author. Fax: +1 865 974 3067.
E-mail addresses: phollin1@utk.edu (P.R. Hollingsworth), chulsey@utk.edu (C. Darrin Hulsey).

1055-7903/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2011.05.020

discordance among gene trees and the desire to infer the best species tree for a group challenges our prior understanding of how phylogenetic relationships should be examined (Maddison, 1997; Maddison and Knowles, 2006; Degnan and Rosenberg, 2006, 2009). For instance, simulations show that simply adding more data to a concatenated matrix can lead to faulty inferences concerning the true species tree when incomplete lineage sorting, or deep coalescence, causes discordance between phylogenetic markers (Kubatko and Degnan, 2007). To account for this shortcoming, methods have been developed that take phylogenetic information from multiple loci and model their evolution under multi-species coalescent theory in an attempt to account for incomplete lineage sorting (Edwards et al., 2007; Kubatko et al., 2009; Liu et al., 2009; Heled and Drummond, 2010). These methods are now being implemented across a broad range of vertebrate groups, but have largely focused on closely related species (White et al., 2009; Carstens and Dewey, 2010; McCormack et al., 2011). Incomplete lineage sorting of gene trees could also be a problem deeper in the phylogenetic history of a clade of organisms, especially if the timeframe for gene coalescence is greater than the period during which groups became genetically isolated. As previous analyses of OPM minnow relationships have had difficulties resolving relationships not only among recently diverged groups but also among genera deep in the clade, variability in gene coalescence may be problematic to phylogenetic inference at several levels in this group.

To evaluate if the stochasticity of coalescence poses a problem for estimating the phylogeny among closely related OPM minnow species, groups such as *Phenacobius* would be interesting to examine within a gene tree/species tree phylogenetic framework. All five named *Phenacobius* species are morphologically diagnosable and generally allopatrically distributed, leaving little question of species monophyly (Etnier and Starnes, 1993; Jenkins and Burkhead, 1994). This lack of sympatry among members of *Phenacobius* should also lead to relatively little hybridization among these five species which is a potentially complicating source of incongruence among gene trees that is not accounted for in most current gene tree/species tree methods (Degnan and Rosenberg, 2009). Additionally, all previous analyses of OPM phylogeny included *Phenacobius* and the placement of this group has varied substantially (Mayden, 1989; Coburn and Cavender, 1992; Simons and Mayden, 1997; Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). Finally, Dimmick and Burr (1999) conducted a study of the phylogenetic relationships among the five species of *Phenacobius* based on a combination of morphological, allozyme, and DNA sequence data which provides a hypothesis with which to compare the results from the molecular phylogenetic approaches taken in this study.

Most previous studies of OPM phylogenetics have relied heavily on mitochondrial and nuclear intron sequence data (Simons and Mayden, 1997; Simons et al., 2003; Mayden et al., 2006). Only recently has sequence data from protein coding nuclear loci been utilized in phylogenetic analyses of NA minnows (Bufalino and Mayden, 2010; Schonhuth and Mayden, 2010; Hulsey and Hollingsworth, 2011). Therefore, a major goal of this study was to generate sequence data for six putatively single copy, protein-coding nuclear loci to serve as phylogenetic markers within NA minnows. In this study, we combine sampling among most OPM genera with complete taxon sampling within the genus *Phenacobius* in order to contrast results from a coalescent-based species tree phylogenetic approach with a concatenated Bayesian analysis of protein-coding nDNA at various levels of phylogenetic relationships.

We used a recently developed multi-species coalescent-based phylogenetic strategy implemented through *BEAST (Heled and Drummond 2010) to examine the relationships among 16 species of OPM minnows, and to compare to results from a concatenated Bayesian analysis. We utilized sequences from seven loci to address a number of questions concerning these relationships. First, we asked what the phylogenetic relationships among exemplars of several of the most diverse OPM genera are. Then, we asked how the genus *Phenacobius* is related phylogenetically to the other OPM minnow genera analyzed in this study. Finally, we examined the phylogenetic relationships among the five species of *Phenacobius*. For each of the above questions, we compared the results between the concatenated Bayesian analysis and the species tree approach implemented in *BEAST (Heled and Drummond, 2010).

2. Materials and methods

2.1. DNA sequence generation

Specimens sequenced in this study were collected in the field using a seine net. Locality information and museum accession numbers are given (Table 1). Specimens of all five species in the OPM genus *Phenacobius* were included, as well as representative taxa from 11 eastern NA OPM genera. These 11 taxa included species designated by Mayden (1989) as chubs (*Campostoma oligolepis*, *Erimystax dissimilis*, *Exoglossum laurae*, and *Nocomis effusus*), as well as taxa designated by Mayden (1989) as shiners (*Cyprinella callistia*, *Luxilus coccogenis*, *Lythrurus fasciolaris*, and *Notropis leuciodus*). We also included the species *Pimephales notatus* and *Hybopsis amblops* that were hypothesized by Mayden (1989) to fall outside of the chub and shiner clades, but that have been recovered as nested within the shiner clade in subsequent phylogenetic analyses (Coburn and Cavender, 1992; Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). *Rhinichthys cataractae* was also included. Morphological phylogenies (Mayden, 1989; Coburn and Cavender, 1992) did not include *Rhinichthys* spp. in the OPM clade. However, more recent molecular phylogenies have included *Rhinichthys* as an early diverging lineage within the OPM radiation and closely related to other chub genera (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). Overall, the taxa included here span the phylogenetic breadth of previously proposed hypotheses of OPM relationships. Individual specimens were anesthetized in MS-222 prior to removal of a pectoral fin for a tissue sample. Tissue samples were stored in 1.5 mL tubes in 95% EtOH and placed in an −80 °C freezer for long-term storage. DNA extraction was performed using a Qiagen DNeasy kit (Qiagen Sciences, MD, USA).

PCR amplification was carried out using an Eppendorf DNA thermocycler. The sequences for all primer sets used in this study are given (Appendix A). The mitochondrial cytochrome *b* (Cytb) gene was amplified using primers from Schmidt and Gold (1993) for all species except *Phenacobius mirabilis*, *Phenacobius teretulus*, and *Phenacobius uranops*. Cytb was amplified for these three species using the primer set MinCytb F2 and MinCytb R1. The six nuclear loci examined included: ectodermal-neural cortex 1 (Enc1), hypothetical protein LOC 564097 (Ptr), novel protein similar to vertebrate ryanodine receptor 3 (Ryr3), protein similar to SH3 and PX domain containing 3 gene (Sh3px3), T-box brain 1 (Tbr1), and zic family member 1 (Zic1). All six nuclear genes were sequenced using primers from Li et al. (2007). PCR conditions consisted of an initial denaturation phase at 94 °C (2 min) followed by 35 cycles of 94 °C (1 min), 54 °C (1 min), and 72 °C (1 min). A final elongation phase of 72 °C (4 min) was performed after the cycles in order to ensure complete elongation of amplified products.

DNA sequencing was performed at the University of Washington's High Throughput Genomics Unit utilizing the same primers used during PCR. Sequence files were contiged using Sequencher 4.8 (Gene Codes, Ann Arbor, MI, USA) and heterozygous sites in

Table 1
Collection locality, latitude/longitude (Lat/Long), and museum accession numbers are given for the voucher specimens utilized in this study. UTEIC = University of Tennessee Etnier Ichthyology Collection, YPM = Yale Peabody Museum.

Species	Locality	Lat/Long	Voucher accession number
Campostoma oligolepis	Little R., TN	35.786N, 83.883W	UTEIC 44.11695
Cyprinella callistia	Conasauga R., TN	34.991N, 84.776W	UTEIC 44.11779
Erimystax dissimilis	Clinch R., TN	36.534N, 83.144W	YPM 19833
Exoglossum laurae	S. Fk. New R., NC	36.221N, 81.640W	UTEIC 44.11728
Hybopsis amblops	Spring Cr., TN	35.224N, 84.512W	UTEIC 44.11793
Luxilus coccogenis	Little R., TN	35.786N, 83.883W	UTEIC 44.11698
Lythrurus fasciolaris	Fishing Cr., KY	37.263N, 84.720W	YPM 19938
Nocomis effusus	Otter Cr., KY	36.666N, 84.973W	YPM 18826
Notropis leuciodus	Little R., TN	35.786N, 83.883W	UTEIC 44.11700
Phenacobius catostomus	Conasauga R., TN	34.991N, 84.776W	UTEIC 44.11786
Phenacobius crassilabrum	N. Toe R., NC	36.024N, 82.023W	UTEIC 44.12060
Phenacobius mirabilis	Pawpaw Cr., TN	36.305N, 89.357W	UTEIC 44.11944
Phenacobius teretulus	S. Fk. New R., NC	36.221N, 81.640W	UTEIC 44.11724
Phenacobius uranops	Chisolm Cr., TN	35.227N, 87.566W	UTEIC 44.12256
Pimephales notatus	Clinch R., TN	36.534N, 83.144W	YPM 19828
Rhinichthys cataractae	S. Fk. New R., NC	36.221N, 81.640W	UTEIC 44.11726
Semotilus atromaculatus	Cox Cr., TN	36.079N, 83.899W	UTEIC 44.11694

the nuclear loci were coded as ambiguous using the IUPAC codes for heterozygous sites. In addition to the newly generated sequences generated in this study, sequence data for Cytb and Enc1 from Hulsey and Hollingsworth (2011) was downloaded from GenBank for several species of NA cyprinids for use in phylogenetic analyses (Table 2). Sequence data for all loci examined were downloaded from GenBank for the Asian species *Danio rerio* which was used as an outgroup in all phylogenetic analyses. Additionally, the non-OPM NA minnow *Semotilus atromaculatus* was included as an outgroup to the OPM minnows. Due to difficulties in PCR amplification, the Cytb sequence data for *S. atromaculatus* was also downloaded from GenBank (Dowling et al., 2002). GenBank accession numbers for all sequence data are given (Table 2). Sequences were aligned using Clustal X (Thompson et al., 1997). Codon sites were defined using MacClade 4.0 (Maddison and Maddison, 2000).

2.2. Gene tree reconstruction

We first inferred the gene tree for each individual locus in order to compare the topologies among the individual genes. Because of the number of informative sites in the mitochondrial gene, each codon position for Cytb was assigned a separate model of molecular evolution. The nuclear loci were not partitioned into individual codon sites because the low variability in their first and second codon positions provided little information to estimate parameters for a separate model of molecular evolution. The first codon position of Zic1 was monomorphic across the North American species analyzed in this study and was therefore excluded in the phylogenetic analyses. The best model of molecular evolution for each locus was chosen using MrModelTest2 (Nylander, 2004). MrModelTest2 starts with a neighbor-joining tree for each partition and then calculates likelihood scores for each of the 24 substitution models that can be implemented in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). The best substitution model for each genetic partition was then chosen based on the model with the lowest AIC score, which penalizes models consisting of more free parameters. Models chosen for each locus are presented along with the length in base pairs of the sequences examined at each locus (Table 3).

Model parameters were then designated in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) in order to approximate the maximum likelihood tree for each locus. For the Cytb analyses, the command prset ratepr = variable was used to allow for rate variation between codon positions. The gamma shape distribution, proportion of invariant sites, state frequencies, and relative rates of substitution were estimated separately for the three codon positions of Cytb using the unlink command in MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). Individual MCMC analyses of each locus consisted of two independent runs of four chains and were run for 1,000,000 generations with trees and parameter estimates sampled every 100 generations. The default heating temperature of 0.1 was used in these analyses. Each MCMC analysis was run three separate times.

Table 2
Genbank accession numbers for new and downloaded sequences used in this study are given.

Species	Cytb	Enc1	Ptr	RYR3	Sh3px3	Tbr1	Zic1
Campostoma oligolepis	HQ446741	HQ446787	JF949831	JF949802	JF949894	JF949863	JF949858
Cyprinella callistia	HQ446743	HQ446768	JF949825	JF949806	JF949882	JF949873	JF949848
Erimystax dissimilis	HQ446746	HQ446798	JF949824	JF949817	JF949891	JF949875	JF949856
Exoglossum laurae	JF949841	JF949840	JF949832	JF949803	JF949892	JF949878	JF949861
Hybopsis amblops	HQ446747	HQ446784	JF949829	JF949809	JF949883	JF949876	JF949849
Luxilus coccogenis	HQ446748	HQ446772	JF949830	JF949811	JF949884	JF949877	JF949846
Lythrurus fasciolaris	HQ446749	HQ446791	JF949827	JF949808	JF949885	JF949871	JF949857
Nocomis effusus	HQ446750	HQ446782	JF949834	JF949805	JF949895	JF949865	JF949859
Notropis leuciodus	HQ446753	HQ446783	JF949828	JF949810	JF949880	JF949874	JF949847
Phenacobius catostomus	HQ446758	HQ446799	JF949821	JF949812	JF949889	JF949866	JF949851
Phenacobius crassilabrum	JF949842	JF949836	JF949819	JF949814	JF949886	JF949867	JF949855
Phenacobius mirabilis	JF949845	JF949839	JF949823	JF949813	JF949890	JF949868	JF949853
Phenacobius teretulus	JF949844	JF949838	JF949822	JF949816	JF949888	JF949870	JF949852
Phenacobius uranops	JF949843	JF949837	JF949820	JF949815	JF949887	JF949869	JF949854
Pimephales notatus	HQ446759	HQ446776	JF949826	JF949807	JF949881	JF949872	JF949850
Rhinichthys cataractae	HQ446760	HQ446779	JF949833	JF949804	JF949893	JF949864	JF949862
Semotilus atromaculatus	AF452082	HQ446781	JF949835	JF949818	JF949896	JF949879	JF949860
Danio rerio	AC024175	EF032975	EF032949	EF032936	EF033001	EF032962	EF032910

Table 3
The length in base pairs (bp) of each partition analyzed is given. Zic1 does not include first position codon sites. Models chosen by MrModelTest2 for each partition (Nylander, 2004) are abbreviated: GTR (general time reversible), HKY (Hasegawa, Kishino, and Yano), K80 (Kimura 80), SYM (symmetrical model), I (proportion of invariant sites), G (gamma distributed substitution rates). Maximum uncorrected sequence divergence is given for each partition, as well as the North American species displaying this divergence. Abbreviated taxa names are detailed in the legend to Fig. 1.

Partition	Length (bp)	Maximum sequence divergence	Model
Cytb 1st	380	11.9% *Sematr – Phecat*	SYM + I + G
Cytb 2nd	380	1.6% *Noceff – Phemir*	HKY
Cytb 3rd	380	52.5% *Sematr – Phemir*	GTR + I + G
Enc1	810	3.8% *Sematr – Eridis*	HKY + G
Ptr	699	2.6% *Sematr – Phecra*	GTR + G
Ryr3	822	3.2% *Sematr – Pimnot/Phecat/Phemir/Pheura*	HKY + I
Sh3px3	705	4.3% *Phecat – Notleu*	K80 + G
Tbr1	645	2.0% *Sematr – Exolau*	HKY + I
Zic1*	572	3.3% *Camoli – Eridis*	HKY + I

Convergence in the MrBayes analyses was assessed by analyzing the split frequencies between the two simultaneous but independent MCMC runs using the "compare" and "cumulative" plots in the program AWTY (Nylander et al., 2008). The program Tracer (Drummond and Rambaut, 2007) was also used to assess convergence by monitoring likelihood and ESS, or effective sample size, values through the course of each MCMC run. The ESS is a proxy for the amount of mixing of Markov chains and represents the number of independent draws from the posterior distribution. High ESS values (>200) signify sufficient mixing of Markov chains, and consequently, low amounts of autocorrelation between parameter estimates from one generation to the next during the course of the MCMC run. Based on all convergence diagnostics, each run had converged by 50,000 generations and the first 100,000 generations for each MCMC search were discarded as the burn-in. Then, the remaining post burn-in trees were used to construct a 50% majority rule consensus tree for each individual locus using the sumt command. Posterior probability values were averaged across the three independent MCMC runs for each locus.

2.3. Multi-locus species tree reconstructions

Two methods were utilized in order to reconstruct the species trees from the combination of the individual loci. For both analyses, we first concatenated the data into a matrix containing 5393 characters. For the concatenated analysis, we used MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003) to generate a concatenated phylogeny using all seven loci. In this analysis, we specified nine partitions corresponding to the three codon positions of Cytb and the six unpartitioned nuclear loci (minus Zic1 first position sites). The same models of molecular evolution that were applied in individual gene tree reconstructions were assigned to the nine partitions (Table 3). The ratepr = variable command was applied in order to allow for rate variation across partitions, and the gamma shape distribution, proportion of invariant sites, state frequencies, and relative rates of substitution were unlinked across partitions. Each concatenated MrBayes run consisted of two separate runs of four chains for 10,000,000 generations with a sampling of trees and parameter values every 1000 generations, and heating temperature of 0.1. Convergence was assessed both in AWTY (Nylander, 2008) and using scale reduction factors reported from MrBayes 3.1.2 (Ronquist and Huelsenbeck, 2003). Tracer (Drummond and Rambaut, 2007) was also used to graphically depict likelihood and ESS values over the course of the runs. Because all runs appeared to have converged by 500,000 generations, the first 1,000,000 runs were subsequently discarded as the burn-in period. We ran three independent MCMC searches for this dataset and averaged the posterior probability values for the nodes across the three replicates to produce our consensus concatenated phylogeny.

The program *BEAST (Heled and Drummond, 2010) was also used to estimate a species tree from the individual loci. First, the partitioned alignment of 5393 characters from the concatenated analysis was imported into the program BEAUti v1.5.4 (Drummond and Rambaut, 2007). We then unlinked substitution models, clock models, and trees across each partition, with the exception of Cytb in which the trees were linked between the three codon positions. The same substitution models used in the concatenated analysis were assigned to each partition (Table 3). A relaxed molecular clock for each partition was estimated relative to Enc1 with all rate estimates drawn from an uncorrelated lognormal distribution. The ploidy level of the Cytb partition was designated as "mitochondrial" and all nuclear loci were designated "autosomal nuclear". Each *BEAST (Heled and Drummond, 2010) search consisted of 1.0×10^8 generations, sampling trees every 1000 generations and used the default priors from BEAUti v1.5.4 (Drummond and Rambaut, 2007). We ran five replicate species tree searches using *BEAST (Heled and Drummond, 2010). Convergence was assessed after importing the log file from each run into Tracer (Drummond and Rambaut, 2007) and then monitoring likelihood and ESS values through the course of the run. To ensure the burn-in was sufficient and to allow our computer to efficiently run the program, the first 9.0×10^7 generations were discarded from each run and then the post burn-in trees and parameter estimates were combined from the five independent runs to produce a majority-rule consensus tree using LogCombiner (Drummond and Rambaut, 2007).

3. Results

3.1. Patterns of sequence variation and individual gene trees

Maximum uncorrected sequence divergence between North American species utilized in this study and the two species displaying the divergence are provided for each partition (Table 3). The most variable partition was Cytb third codon position (52.5%) and the least was Tbr1 (2.0%). The divergence reported for the Zic1 locus (Table 3) does not include first position codon sites, as this nucleotide position was invariable. The non-OPM creek chub, *S. atromaculatus*, was one of the species involved in the maximum observed sequence divergence in six of the nine partitions with the exception of the Cytb second position, Sh3px3, and Zic1 (Table 3).

The Bayesian 50% majority-rule consensus tree for each individual locus is depicted (Fig. 1). Relationships among the included genera of OPM minnows varied between individual gene trees. The non-OPM taxon *S. atromaculatus* was resolved as falling outside of, or within a polytomy that was sister to, all other NA lineages sampled in five of the seven individual gene trees. In the Sh3px3 gene tree, *S. atromaculatus* was recovered as sister to *N. effusus* with low posterior support (56%). In the Ptr phylogeny, *S. atromaculatus* and *N. effusus* were recovered as more closely related to *Erimystax* + *Phenacobius* spp. than to other NA lineages sampled. This relationship received moderate posterior support (81%) in this gene tree.

The remaining chub lineages, *C. oligolepis*, *E. laurae*, and *R. cataractae*, were resolved as diverging early in the majority of the individual gene trees (4 of 7) along with *N. effusus* (Fig. 1). However, relationships between these taxa and the remainder of the OPM taxa included in this study varied considerably, often receiving low posterior support, depending on the individual gene tree. For example, *E. laurae* is strongly supported (91% posterior) as diverging after *N. effusus*, but before the remainder of the OPM minnows in the Cytb gene tree. However, in the Enc1 gene tree

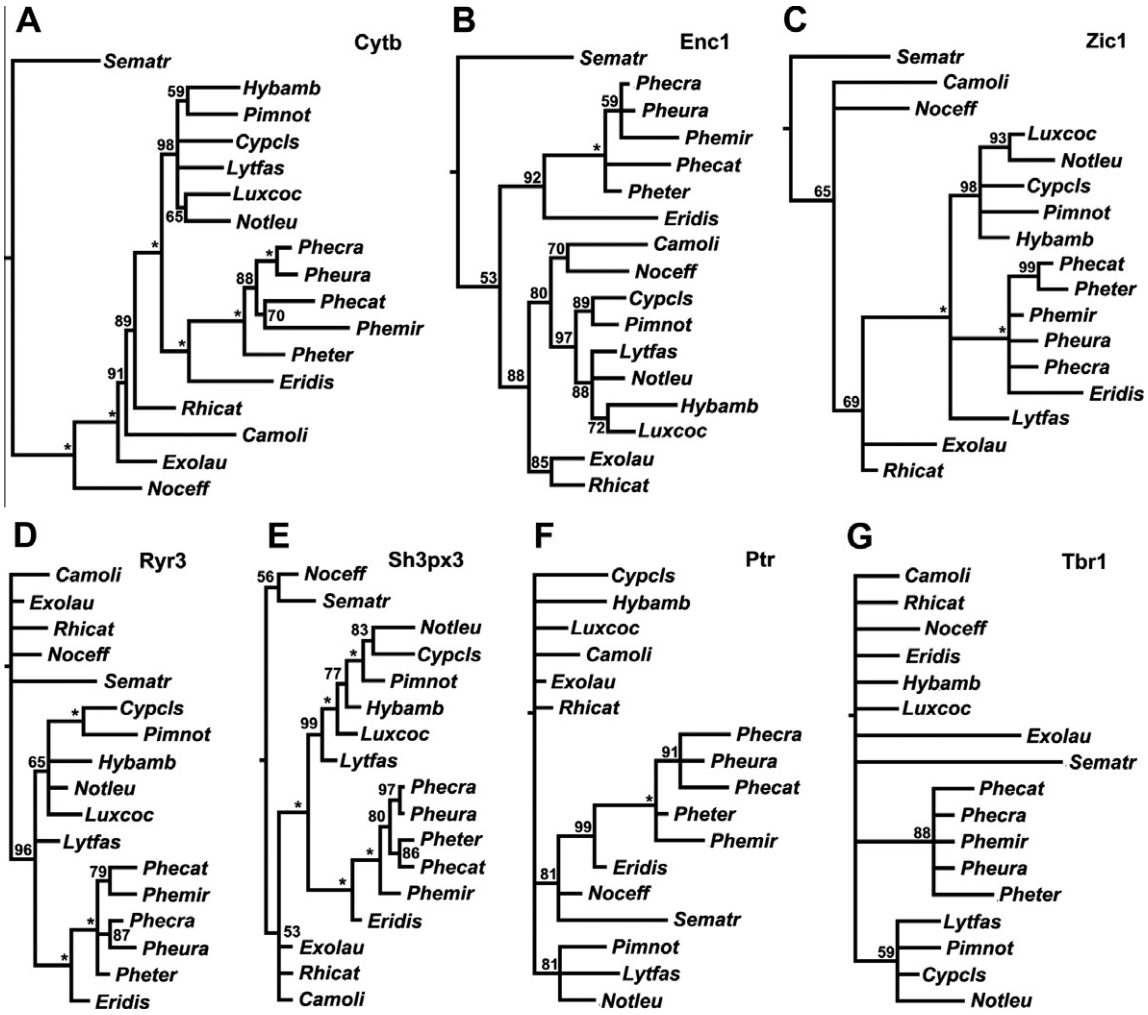

Fig. 1. The 50% majority rule consensus gene trees obtained using MrBayes are presented for the individual loci: (A) Cytb. (B) Enc1. (C) Zic1. (D) Ryr3. (E) Sh3px3. (F) Ptr. (G) Tbr1. Values at nodes are Bayesian posterior support values given in percentages. Asterisks represent 100% posterior support. Taxon names have been abbreviated to save space. *Campostoma oligolepis* (Camoli), *Cyprinella callistia* (Cypcls), *Erimystax dissimilis* (Eridis), *Exoglossum laurae* (Exolau), *Hybopsis amblops* (Hybamb), *Luxilus coccogenis* (Luxcoc), *Lythrurus fasciolaris* (Lytfas), *Nocomis effusus* (Noceff), *Notropis leuciodus* (Notleu), *Phenacobius catostomus* (Phecat), *P. crassilabrum* (Phecra), *P. mirabilis* (Phemir), *P. teretulus* (Pheter), *P. uranops* (Pheura), *Pimephales notatus* (Pimnot), *Rhinichthys cataractae* (Rhicat), *Semotilus atromaculatus* (Sematr). *Danio rerio* was used to root the phylogenies and this branch was subsequently removed from each gene tree for presentation in this figure.

E. laurae and *R. cataractae* are moderately supported as sister taxa (85% posterior). Relationships between the shiner taxa included in this study were also variable. Only two sets of relationships between shiner taxa were resolved in more than one individual gene tree. The sister relationship between *P. notatus* and *C. callistia* was recovered with 100% and 89% posterior support in the Ryr3 and Enc1 gene trees respectively. *L. coccogenis* and *N. leuciodus* were also recovered as each other's closest relative with 93% and 65% posterior support in the Zic1 and Cytb gene trees respectively (Fig. 1).

E. dissimilis was recovered as sister to *Phenacobius* spp. with high posterior support (>90%) in all individual gene trees except for the generally poorly resolved Tbr1 gene tree (Fig. 1). This clade of *Erimystax* + *Phenacobius* was recovered as closely aligned with the NA shiners (*C. callistia*, *H. amblops*, *L. fasciolaris*, *L. coccogenis*, and *N. leuciodus*) in four of the seven individual gene trees with substantial posterior Bayesian support (>95%). Relationships between *Erimystax* + *Phenacobius* and the other OPM lineages sampled were poorly resolved in the Ptr and Tbr1 gene trees. However, in the Enc1 gene tree *Erimystax* + *Phenacobius* was recovered in 53% of post burn-in trees as sister to a moderately supported clade (88% posterior) of all remaining OPM minnows (Fig. 1).

Phenacobius was significantly supported as monophyletic in 6 of 7 gene trees (Fig. 1). However, there was little agreement among the gene trees concerning the phylogenetic relationships within *Phenacobius*, and many, but not all, of these relationships received low posterior support values. The most consistently resolved node within *Phenacobius* was a sister relationship between *P. uranops* and *P. crassilabrum*. This relationship was recovered in 87%, 97%, and 100% of the post burn-in Ryr3, Seh3px3, and Cytb gene trees, respectively. Relationships between the three other members of *Phenacobius* varied considerably across gene trees. *Phenacobius catostomus* and *P. mirabilis* were recovered as sister lineages with moderate posterior support, 70% and 79%, in the Cytb and Ryr3 gene trees (Fig. 1). The two species *P. teretulus* and *P. catostomus* were recovered as sister species in the Sh3px3 and Zic1 gene trees with 86% and 99% posterior probability values respectively.

3.2. Multi-locus species trees

In the concatenated analysis, *S. atromaculatus* was recovered as the outgroup to a strongly supported (100% posterior) clade consisting of the remaining NA lineages sampled (Fig. 2A). All but four nodes in the concatenated topology received significant posterior

support of >95%. These remaining ambiguous nodes subtended the following sets of taxa: (1) *P. mirabilis* and *P. uranops* (54%), (2) *P. catostomus* and *P. teretulus* (82%), (3) *H. amblops* and *R. cataractae* (89%), and (4) *E. laurae* and *R. cataractae* (88%). The chub genera were recovered at the base of the concatenated OPM phylogeny with *N. effusus* strongly supported (100% posterior) as the earliest diverging lineage of the OPM minnows analyzed. This divergence was followed by divergence of other chub genera, first *Campostoma* and then a clade of *Exoglossum* + *Rhinichthys*, with moderate to strong posterior support (88–100%) (Fig. 2A). Within the shiners, *L. fasciolaris* was recovered as the earliest diverging lineage. *H. amblops* and *L. coccogenis* formed a strongly supported clade, while *N. leuciodus* was recovered as sister to a clade of *P. notatus* + *C. callistia*. All relationships within the shiners received high (98–100%) Bayesian posterior support in the concatenated analysis (Fig. 2A).

Phenacobius was strongly supported (100% posterior) as monophyletic and *E. dissimilis* was strongly supported as sister to *Phenacobius* spp. (100% posterior) in the concatenated phylogeny (Fig. 2A). The clade containing both *Erimystax* + *Phenacobius* received 100% posterior support as being sister to the shiners included in this analysis. Within *Phenacobius*, a weakly supported (54% posterior) clade of *P. mirabilis* and *P. uranops* + *P. crassilabrum* was recovered as sister to *P. catostomus* + *P. teretulus* (Fig. 2A).

The topology estimated by *BEAST (Heled and Drummond, 2010) is very similar to the concatenated topology (Fig. 2B). However, posterior support values were noticeably lower in this coalescent-based phylogenetic reconstruction. Therefore, all support values including those below 50% on the *BEAST topology are presented as maximum clade credibility scores in order to compare with results from the concatenated topology. Using the *BEAST (Heled and Drummond, 2010) species tree approach, the chubs are recovered as diverging first from the remaining OPM lineages followed by divergence of *Erimystax* + *Phenacobius* from the shiner clade (Fig. 2B). However, the species tree analysis differed from the concatenated analysis in specific relationships both among the chubs and shiners (Fig. 2). *N. effusus* was again recovered as the earliest diverging OPM lineage, albeit with low posterior support (51%). However, *C. oligolepis* was strongly supported (94% posterior) as sister to the *Exoglossum* + *Rhinichthys* clade contrary to the concatenated analysis. Within the shiners, *L. fasciolaris* was recovered as sister to *N. leuciodus* with low posterior support (37%). Among the shiners, only the *P. notatus* + *C. callistia* clade resolved in the concatenated analysis was again recovered with high posterior support (91%) by the *BEAST (Heled and Drummond, 2010) species tree analysis (Fig. 2B).

E. dissimilis was strongly supported (100% posterior) as sister to a strongly supported (100% posterior) monophyletic *Phenacobius* in the coalescent-based species tree (Fig. 2B). This clade was recovered in 100% of post burn-in species trees as more closely related to the shiner taxa utilized. However, within *Phenacobius*, a different set of relationships was resolved by *BEAST (Heled and Drummond, 2010) than those recovered in the concatenated analysis (Fig. 2). Although the two species *P. uranops* and *P. crassilabrum* were again recovered as sister lineages, this clade was recovered as sister to a clade of *P. catostomus* + *P. teretulus* in 46% of the post burn-in species trees, with *P. mirabilis* resolved as the earliest diverging member of *Phenacobius* (Fig. 2B).

4. Discussion

The individual gene trees of eastern NA OPM minnows were highly discordant. The two species tree approaches also produced phylogenetic hypotheses that differed topologically in several areas. Additionally, our concatenated phylogeny recovered much higher posterior support values as compared to those obtained from the coalescent-based species tree strategy. Despite incongruence between the two multi-locus approaches, both methods of reconstructing the species tree provided a generally concordant backbone topology. Both phylogenies indicated that OPM shiners are monophyletic and nested within the OPM clade, making chubs a paraphyletic group. Both species trees also resolved *Phenacobius* + *Erimystax* as more closely related to the shiner taxa than to the chub taxa that we analyzed. The topologies of the two multi-locus trees are likely to have been strongly influenced by the phylogenetic signal within the Cytb locus due to its greater variability relative to the nuclear loci at this phylogenetic level. In the future, it will be interesting to see if sequencing many more nuclear genes

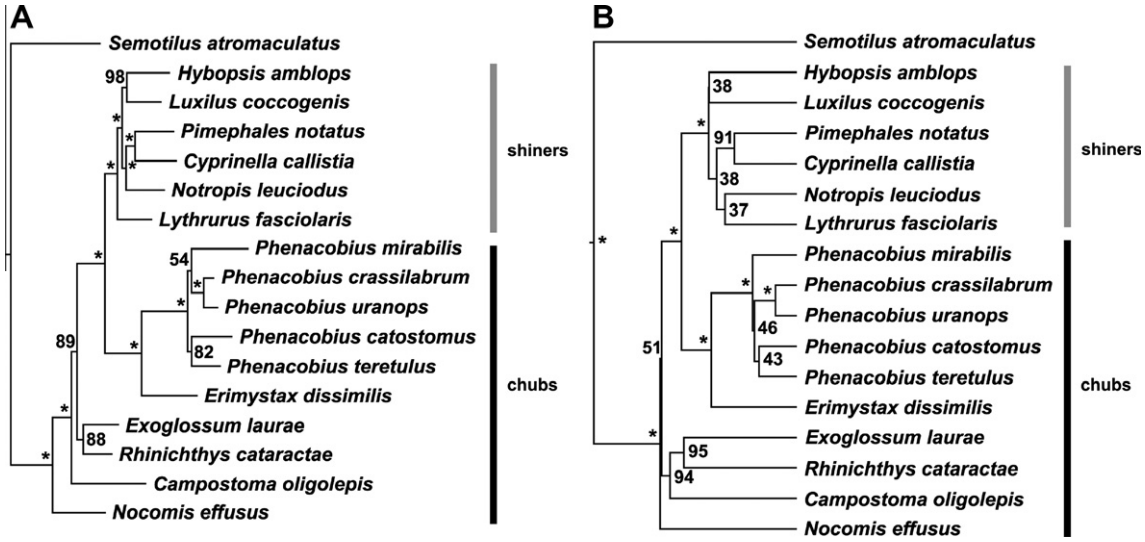

Fig. 2. (A) The 50% majority rule consensus phylogeny is presented from the concatenated Mr. Bayes analysis. (B) The maximum clade credibility phylogeny is presented from the *BEAST species tree analysis. Values at nodes are posteriors given in percentages. Asterisks indicate 100% of post burn-in trees contained the clade. "Chub" and "shiner" designations are based on the morphological phylogenetic hypothesis of Mayden (1989). *Danio rerio* was used to root the phylogenies and this branch was subsequently removed from each species tree for presentation in this figure.

provides a topology that is consistently divergent from that of Cytb. We will also be able to determine if these nuclear genes converge on a singular set of relationships for regions of the OPM minnow phylogeny that are currently difficult to resolve.

One area of discordance between the two multi-locus analyses involved relationships among the chub genera. For example, relationships among the early diverging chub lineages were strongly supported in the concatenated species tree, with *Nocomis* recovered as the earliest diverging OPM taxa. In the coalescent species tree, this sister relationship between *Nocomis* and the remainder of the OPM minnows was not well supported. Additionally, *Campostoma* was recovered as the next chub lineage to diverge after *Nocomis* from the remaining OPM lineages in the concatenated analysis. However, in the coalescent approach *Campostoma* is supported as sister to a clade of *Exoglossum* + *Rhinichthys*. These sets of relationships among chub lineages differed from all previous, large-scale molecular phylogenies of OPM minnows (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011), in which *Nocomis* and *Campostoma* were hypothesized to be sister lineages. Interestingly, the sister group relationship between *Exoglossum* and *Rhinichthys* that was supported by both of our species tree analyses was also recovered by Simons et al. (2003). Mayden et al. (2006) and Hulsey and Hollingsworth (2011) did not include *Exoglossum* species in their analyses, but this relationship was not recovered in the phylogeny presented by Bufalino and Mayden (2010).

Relationships and posterior support values also differed substantially between our two multi-locus approaches within the shiner clade. The concatenated analysis produced a fully resolved, and well-supported, topology of shiner genera relationships. However, the coalescent-based species tree generally provided little support for shiner intergeneric relationships. The only significantly supported node within the shiner clade that we recovered using both species tree strategies was the sister relationship between *Pimephales* and *Cyprinella*. Hulsey and Hollingsworth (2011) presented a topology containing *Pimephales* nested within an unsupported, yet monophyletic *Cyprinella*. However, this relationship was not recovered by any of the other aforementioned studies of OPM molecular phylogenies (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010).

The lower posterior probability values on the coalescent-based species tree can be partially explained by the more parameter-rich phylogenetic model implemented by this strategy. The increased number of parameters in the coalescent model results from taking into consideration the discordance between individual gene trees. Coalescent methods must also estimate parameters affecting coalescent processes such as ancestral population sizes (Heled and Drummond, 2010). Concatenation of data ignores this discordance between gene trees, estimates less parameters, and can lead to inflated posterior support values (Kubatko and Degnan, 2007). The discordance we recovered between our individual gene trees suggest that accounting for the coalescent process is likely warranted when examining relationships among OPM genera and at multiple other levels of phylogenetic inference.

Despite incongruence and variable support between our two multi-locus phylogenetic approaches in several areas of the phylogeny, both analyses recovered the strongly supported, sister relationship between *Phenacobius* + *Erimystax* and the shiner clade. This relationship is consistent with other recent molecular phylogenies produced for this group of fishes (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). However, earlier morphological and molecular phylogenies proposed that *Erimystax* and *Phenacobius* were more closely related to the remaining chub OPM lineages sampled in this study (Mayden, 1989; Coburn and Cavender, 1992; Simons and Mayden, 1997). *Erimystax* and *Phenacobius* are benthic taxa and generally chub-like in appearance. As such, our species tree analyses suggest that chub-like morphology and its association with benthic habitats were common early in the history of the eastern NA OPM clade and persisted at least until the divergence of *Phenacobius* + *Erimystax* from the remaining eastern OPM minnows. Interestingly, the number of species in these early diverging chub lineages is relatively low in comparison to extant shiner diversity. The pelagic shiners likely arose from ancestral benthic forms and subsequently diversified to generate a substantial portion of contemporary OPM diversity. Evolution along a benthic/limnetic habitat continuum has been demonstrated to be a common axis of diversification between closely related species in several lacustrine fish groups (Hatfield and Schluter, 1999; Barluenga et al., 2006; Bertrand et al. 2008). This study suggests that this benthic/limnetic axis of diversification could also be an important macroevolutionary force in generating species diversity across broader phylogenetic scales within lotic environments such as the streams of eastern NA.

There was little agreement between gene trees and species trees in resolving the phylogenetic relationships within *Phenacobius*. However, both of our multi-locus phylogenetic analyses resolved different relationships than those posited by Dimmick and Burr (1999). Their analysis based on morphological and genetic data recovered *P. mirabilis* as the earliest diverging lineage within *Phenacobius*, followed in order of divergence by *P. teretulus*, then by *P. catostomus*, and finally by a clade of *P. uranops* + *P. crassilabrum*. Both our concatenated and species-tree analyses also supported the sister species relationship between *P. uranops* and *P. crassilabrum*, the two species with abutting allopatric ranges in the upper Tennessee River system. However, contrary to the hypothesis of Dimmick and Burr (1999), the species *P. teretulus* and *P. catostomus* were weakly supported as each other's closest relative in both of our species tree approaches. Furthermore, our concatenated analysis recovered a weakly supported sister relationship between *P. mirabilis* and the clade of *P. crassilabrum* + *P. uranops*. A lack of phylogenetically informative variation within *Phenacobius* at the nDNA loci utilized could partially explain the incongruence between gene trees and uncertainty in our species trees.

Several of the same regions of the phylogeny that we had trouble resolving in this study (relationships between chubs lineages and within the shiners and *Phenacobius*) have differed topologically in previous molecular phylogenetic studies of OPM relationships (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). Furthermore, these regions have consistently been recovered as areas with particularly short branch lengths between diverging lineages (Simons et al., 2003; Mayden et al., 2006; Bufalino and Mayden, 2010; Hulsey and Hollingsworth, 2011). These are regions of the phylogeny that have likely experienced large amounts of incomplete lineage sorting of slowly evolving genetic markers due to short time frames between diversification events. Future studies of OPM minnow phylogenetics should aim at generating sequence data from faster evolving nuclear loci, and employ coalescent-based species tree methods, in order to tease apart relationships in these problematic regions of the OPM phylogeny.

With the increasing availability of genomic data from across the tree of life, the ability to analyze DNA sequence data for multiple loci will no longer be limiting to phylogenetics. A transition from single locus to multi-locus, coalescent-based phylogenetics is well underway (Maddison and Knowles, 2006; Edwards et al., 2007; Degnan and Rosenberg, 2009; Kubatko et al., 2009; Liu et al., 2009; Heled and Drummond, 2010; Hulsey et al., 2011). Only by gathering data from independent genetic regions across the genomes of organisms, may we gain a better understanding of the phylogenetic relationships recorded in each locus. Individual gene

phylogenies may then be analyzed separately under coalescent-based methods in order to approximate the distribution of species trees suggested by individual locus data. Future studies of phylogenetics of the hyper-diverse North American OPM minnows should give less weight to phylogenetic hypotheses based on single genetic partitions and concatenated multi-locus matrices and place more emphasis on producing phylogenies that explicitly model the coalescent process.

Acknowledgments

We thank Tom Near and the Yale Peabody Museum of Natural History for providing tissue samples for some taxa utilized in this study. We also thank the University of Tennessee Etnier Ichthyology Collection for allowing us to deposit individual voucher specimens. Ben Keck kindly provided helpful comments on earlier versions of this manuscript. The University of Tennessee provided funding for this research.

Appendix A. Primer names and sequences utilized in this study are given

Locus	Primer name	Sequence
Cytb	H15915	5′-CAACGATCTCCGGTTTACAAGAC-3’
	L14724	5′-GTGACTTGAAAAACCACCGTTG-3′
	MinCytb F2	5′-CGTTGTAATTCAACTACAGGAAC-3′
	MinCytb R1	5′-TTCGGATTACAAGACCGATGCT-3′
Zic1	F9	5′-GGACGCAGGACCGCARTAYC-3′
	R967	5′-CTGTGTGTGTCCTTTTGTGRATYTT-3′
Ryr3	F22	5′-TCGGTAAGCARATGGTGGACA-3′
	R931	5′-AGAATCCRGTGAAGAGCATCCA-3′
Ptr	F458	5′-AGAATGGATWACCAACACYTACG-3′
	R1248	5′-TAAGGCACAGGATTGAGATGCT-3′
Tbr1	F1	5′-TGTCTACACAGGCTGCGACAT-3′
	R820	5′-GATGTCCTTRGWGCAGTTTTT-3′
Enc1	F85	5′-GACATGCTGGAGTTTCAGGA-3′
	R982	5′-ACTTGTTRGCMACTGGGTCAAA-3′
Sh3px3	F532	5′-GACGTTCCCATGATGGCWAAAAT-3′
	R1299	5′-CATCTCYCCGATGTTCTCGTA-3′

References

Barluenga, M., Stolting, K.N., Salzburger, W., Muschick, M., Meyer, A., 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. Nature 439, 719–723.

Bertrand, M., Marcogliese, D.J., Magnan, P., 2008. Trophic polymorphism in brook charr revealed by diet, parasites and morphometrics. J. Fish Biol. 72, 555–572.

Bufalino, A.P., Mayden, R.L., 2010. Molecular phylogenetics of North American phoxinins (Actinopterygii: Cypriniformes: Leuciscidae) based on RAG1 and S7 nuclear DNA sequence data. Mol. Phylogenet. Evol. 55, 274–283.

Burr, B.M., Mayden, R.L., 1992. Phylogenetics and North American freshwater fishes. In: Mayden, R.L. (Ed.), Systematics, Historical Ecology, and North American Freshwater Fishes. Stanford University Press, Stanford, CA, pp. 18–75.

Carstens, B.C., Dewey, T.A., 2010. Species delimitation using a combined coalescent and information – theoretic approach: an example from North American Myotis bats. Syst. Biol. 59, 400–414.

Coburn, M.M., Cavender, T.M., 1992. Interrelationships of North American freshwater fishes. In: Mayden, R.L. (Ed.), Systematics, Historical Ecology, and North American Freshwater Fishes. Stanford University Press, Stanford, CA, pp. 328–373.

Degnan, J.H., Rosenberg, N.A., 2006. Discordance of species trees with their most likely gene trees. PLoS Genet. 2 (5), e68.

Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24, 332–340.

Dimmick, W.W., Burr, B.M., 1999. Phylogenetic relationships of the suckermouth minnows, genus Phenacobius, inferred from parsimony analyses of nucleotide sequence, allozymic and morphological data (Cyprinidae: Cypriniformes). Biochem. Syst. Ecol. 27, 469–485.

Dowling, T.E., Tibbets, C.A., Minckley, W.L., Smith, G.R., 2002. Evolutionary relationships of the plagopterins (Teleostei: Cyprinidae) from cytochrome b sequences. Copeia 2002, 665–678.

Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol. Biol. 7, 214.

Edwards, S.V., Liu, L., Pearl, D.K., 2007. High-resolution species trees without concatenation. Proc. Natl. Acad. Sci. USA 104, 5936–5941.

Etnier, D.A., Starnes, W.C., 1993. The Fishes of Tennessee. The University of Tennessee Press, Knoxville, TN.

Hatfield, T., Schluter, D., 1999. Ecological speciation in sticklebacks: environment-dependent hybrid fitness. Evolution 53, 866–873.

Heled, J., Drummond, A.J., 2010. Bayesian inference of species trees from multilocus data. Mol. Biol. Evol. 27, 570–580.

Hulsey, C.D., Keck, B.P., Hollingsworth Jr., P.R., 2011. Species tree estimation and the historical biogeography of heronie cichlids. Mol. Phylogenet. Evol. 58, 124–131.

Hulsey, C.D., Hollingsworth Jr., P.R., 2011. Do constructional constraints influence cyprinid (Cyprinidae: Leuciscinae) craniofacial coevolution? Biol. J. Linn. Soc. 103, 136–146.

Jenkins, R.E., Burkhead, N.M., 1994. Freshwater Fishes of Virginia. American Fisheries Society, Bethesda, MD.

Kubatko, L.S., Degnan, J.H., 2007. Inconsistency of phylogenetic estimates from concatenated data under coalescence. Syst. Biol. 56, 17–24.

Kubatko, L.S., Carstens, B.C., Knowles, L.L., 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25, 971–973.

Li, C.H., Orti, G., Zhang, G., Lu, G.Q., 2007. A practical approach to phylogenomics: the phylogeny of ray-finned fish (Actinopterygii) as a case study. BMC Evol. Biol. 7.

Liu, L., Yu, L., Kubatko, L., Pearl, D.K., Edwards, S.V., 2009. Coalescent methods for estimating phylogenetic trees. Mol. Phylogenet. Evol. 53, 320–328.

Maddison, W.P., 1997. Gene trees in species trees. Syst. Biol. 46, 523–536.

Maddison, D.R., Maddison, W.P., 2000. Macclade. Sinauer Associates, Inc., Sunderland, MA.

Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55, 21–30.

Mayden, R.L., 1989. Phylogenetic studies of North American minnows, with emphasis on the genus Cyprinella (Teleostei: Cypriniformes). Misc. Pub. Univ. Kans. Mus. Nat. Hist 80, 1–189.

Mayden, R.L., Simons, A.M., Wood, R.M., Harris, P.M., Kuhajda, B.R., 2006. Molecular systematics and classification of North American Notropin shiners and minnows (Cypriniformes: Cyprinidae). In: De Lourdes Lozano-Villano, M.A., Contreras-Balderas, A.J. (Eds.), Studies of North American Desert Fishes in Honor of E.P. (Phil) Pister, Conservationist. Universidad Autonoma de Nuevo Leon, Nuevo Leon, Mexico, pp. 72–101.

McCormack, J.E., Heled, J., Delaney, K.S., Townsend Peterson, A., Knowles, L.L., 2011. Calibrating divergence times on species trees versus gene trees: implication for speciation history of Aphelocoma jays. Evolution 65, 184–202.

Nylander, J.A.A., 2004. MrModelTest v2. Uppsala University.

Nylander, J.A.A., Wilgenbusch, J.C., Warren, D.L., Swofford, D.L., 2008. AWTY (are we there yet?): a system for graphical exploration of MCMC convergence in Bayesian phylogenetics. Bioinformatics 24, 581–583.

Ronquist, F., Huelsenbeck, J.P., 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19, 1572–1574.

Schmidt, T.R., Gold, J.R., 1993. Complete sequence of the mitochondrial cytochome-b gene in the cherryfin shiner, *Lythrurus roseipinnis* (Teleostei, Cyprinidae). Copeia 1993, 880–883.

Schonhuth, S., Mayden, R.L., 2010. Phylogenetic relationships of the genus Cyprinella based on mitochondrial and nuclear gene sequences. Mol. Phylogenet. Evol. 55, 77–98.

Simons, A.M., Mayden, R.L., 1997. Phylogenetic relationships of the creek chubs and the spine-fins: an enigmatic group of North American cyprinid fishes (Actinopterygii: Cyprinidae). Cladistics 13, 187–205.

Simons, A.M., Berendzen, P.B., Mayden, R.L., 2003. Molecular systematics of North American phoxinin genera (Actinopterygii: Cyprinidae) inferred from mitochondrial 12S and 16S ribosomal RNA sequences. Zool. J. Linn. Soc. – Lond. 139, 63–80.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., Higgins, D.G., 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucl. Acids Res. 25, 4876–4882.

White, M.A., Ane, C., Dewey, C.N., Larget, B.R., Payseur, B.A., 2009. Fine-scale phylogenetic discordance across the house mouse genome. PLoS Genet. 5 (11), e1000729.